



TITLE:

Face and eye scanning in gorillas (*Gorilla gorilla*), orangutans (*Pongo abelii*), and humans (*Homo sapiens*): unique eye-viewing patterns in humans among hominids.

AUTHOR(S):

Kano, Fumihiro; Call, Josep; Tomonaga, Masaki

CITATION:

Kano, Fumihiro ...[et al]. Face and eye scanning in gorillas (*Gorilla gorilla*), orangutans (*Pongo abelii*), and humans (*Homo sapiens*): unique eye-viewing patterns in humans among hominids.. *Journal of comparative psychology* 2012, 126(4): 388-398

ISSUE DATE:

2012-11

URL:

<http://hdl.handle.net/2433/172546>

RIGHT:

© 2012 American Psychological Association; この論文は著者最終稿です。内容が印刷版と異なることがありますので、引用の際には出版社版をご確認ご利用ください。This is the Accepted Author Manuscript. Please cite only the published version.

Research article

Face and eye scanning in gorillas (*Gorilla gorilla*), orangutans (*Pongo abelii*), and humans (*Homo sapiens*): unique eye-viewing patterns in humans among hominids

Fumihiro Kano (1, 2)

Josep Call (3)

Masaki Tomonaga (1)

1. Primate Research Institute, Kyoto University, Inuyama, Japan

2. Japan Society for Promotion of Science

3. Max Plank Institute for Evolutionary Anthropology, Leipzig, Germany

fkanou@pri.kyoto-u.ac.jp

+81-80-6902-5013

This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

19 Abstract

20 Because the faces and eyes of primates convey a rich array of social information, the way in which
21 primates view faces and eyes reflects species-specific strategies for facial communication. How are
22 humans and closely related species such as great apes similar and different in their viewing patterns for
23 faces and eyes? Following previous studies comparing chimpanzees with humans, this study used the
24 eye-tracking method to directly compare the patterns of face and eye scanning by humans, gorillas,
25 and orangutans. Human and ape participants freely viewed pictures of whole bodies and full faces of
26 conspecifics and allospecifics under the same experimental conditions. All species were strikingly
27 similar in that they viewed predominantly faces and eyes. No particular difference was identified
28 between gorillas and orangutans, and they also did not differ from the chimpanzees tested in previous
29 studies. However, humans were somewhat different from apes, especially with respect to prolonged
30 eye viewing. We also examined how species-specific facial morphologies, such as the male flange of
31 orangutans and the black–white contrast of human eyes, affected viewing patterns. Whereas the male
32 flange of orangutans affected viewing patterns, the color contrast of human eyes did not. Humans
33 showed prolonged eye viewing independently of the eye color of presented faces, indicating that this
34 pattern is internally driven rather than stimulus dependent. Overall, the results show general
35 similarities among the species and also identify unique eye-viewing patterns in humans.

36

37 Introduction

38 Human and non-human primates have highly sophisticated forms of facial communication. Faces and
39 eyes convey a wide variety of information such as identity, age, sex, emotion, and direction of
40 attention. Thus, primates perceive faces and eyes differently from the way they perceive other visual
41 stimuli (Emery, 2000; Tomonaga, 2010). Moreover, primates often engage in prolonged face-to-face
42 interactions that are accompanied by eye contact (looking into another individual's eyes) or its
43 avoidance (Gomez, 1996; Kleinke, 1986; Thomsen, 1974). How and when such unique forms of facial
44 communication evolved in primates have long been of interest, particularly from a comparative
45 perspective (Chevalier-Skolnikoff, 1973; Darwin, 1872/1999; van Hooff, 1967).

46 Humans and great apes such as chimpanzees, gorillas, and orangutans are known to share
47 several forms of facial communication with one another. For example, prolonged eye contact indicates
48 mild threat, and thus gaze avoidance indicates submission (Goodenough, McGuire, & Jakob, 1993;
49 Redican, 1975). Eye contact is also used in an affiliative context, such as the solicitation of play/sex
50 and post-conflict appeasement (de Waal, 1990; Goodall, 1986; Kano, 1980; Yamagiwa, 1992). Gaze
51 following (looking in the same direction) frequently occurs after the establishment of eye contact with
52 another individual (Bräuer, Call, & Tomasello, 2005; Itakura, 2004; Okamoto-Barth, Call, &
53 Tomasello, 2007; Tomasello, Hare, Lehmann, & Call, 2007). Additionally, studies have suggested that
54 these species use eye contact ostensively (i.e., viewing another individual's eyes with communicative
55 intent) (Gomez, 1996).

56 On the other hand, notable differences among the species have also been identified. For
57 example, although close-range long-bout affiliative communication typically accompanies intense eye
58 contact in humans, this kind of communication seems to occur more frequently in tactile than in visual
59 forms in great apes (e.g., grooming; Goodall, 1986; Shaller, 1963). Several studies have suggested that
60 mutual gaze (returning another individual's gaze) is uncommon and that gaze avoidance is frequent in

gorillas (Shaller, 1963; but see Yamagiwa, 1992) and orangutans (Kaplan & Rogers, 2002). Kobayashi and Kohshima (2001) found that, compared with other primates, humans have exceptionally large white sclera that clearly contrasts with the colors of their iris and skin. These authors hypothesized that human eyes have evolved to enhance gaze signals such as eye direction. Tomasello et al. (2007) found that great apes (chimpanzees, bonobos, and gorillas) were less sensitive than human infants were to the eye direction of a human experimenter when the experimenter's head was immobile. Okamoto-Barth et al. (2007) found that, compared with chimpanzees, bonobos, and gorillas, orangutans were less sensitive to gaze (head) directions and more attracted to target objects.

These results suggest that, although humans and great apes share a basic set of facial communicative skills, significant species differences characterize their habitual styles. Direct comparisons are necessary to reveal such differences among these species. Kano and Tomonaga (2009, 2010) used the eye-tracking method to measure how humans and chimpanzees scanned the faces and eyes of conspecifics and allospecifics. Both species were presented with pictures of whole bodies and faces, and they viewed the pictures freely. The species were strikingly similar in their patterns of face and eye scanning. For example, both species fixated on faces and eyes more frequently than on other parts of bodies and faces. However, several species differences were also identified. For example, chimpanzees viewed faces and eyes more briefly than did humans; typically, chimpanzees only glanced at eyes, whereas humans viewed both eyes (left and right eyes) alternately. Additionally, when presented with facial expressions including conspicuous mouth actions, humans viewed the eyes rather than the mouth, and chimpanzees viewed the mouth rather than the eyes. Thus, previous studies have shown both striking similarities among the species and unique eye-viewing patterns in humans.

At present, such face- and eye-scanning data are not available for other ape species. Thus, this study aimed to obtain a broader comparative picture about this issue. Following previous studies comparing chimpanzees with humans ([Kano & Tomonaga, 2009](#), [2010](#)), we compared the patterns of

face and eye scanning by humans, gorillas, and orangutans, We had three primary reasons for examining the patterns of face and eye scanning in gorillas and orangutans. First, gorillas and especially orangutans are phylogenetically more distant from humans than are chimpanzees. Thus, this study enabled us to test whether the unique eye-viewing patterns of humans indicate general differences between humans and great apes or isolated differences among species. Based on previous studies showing intense eye contact, especially in humans, it would be expected that humans would view eyes for longer durations than do gorillas and orangutans. Second, several of the previous studies have reported a high probability of gaze avoidance in gorillas and orangutans (Kaplan & Rogers, 2002; Shaller, 1963). Additionally, orangutans lead semi-solitary lives in their natural habitats, which is the least socially specialized form among the great apes. The simple experimental design of this study (presenting facial pictures) enabled us to test gorillas' and orangutans' default motivation for viewing eyes (i.e., in the absence of social interaction).

Finally, each species has species-specific morphological facial features that have presumably evolved for purposes of communication. This study enabled us to test the role of these facial features in visual perception. For example, humans have a color contrast between the dark iris and white sclera in their eyes (Kobayashi & Kohshima, 2001), adult male gorillas have a high crest on the top of their heads, and adult male orangutans have developed flanges on the sides that are thought to be sexual signals (Ankel-Simons, 2000; Kuze, Malim, & Kohshima, 2005). Infant and juvenile orangutans have pale coloring around their eyes and mouth, thought to signal immaturity (Kuze, et al., 2005). As those facial features are visually conspicuous, they would be expected to automatically attract viewers' attention. That is, the viewing patterns would be expected to be dependent on the presence/absence of such conspicuous facial features. However, if viewing patterns were independent of the types of faces presented, such patterns would be internally driven rather than stimulus dependent.

109 Following previous studies comparing chimpanzees with humans ([Kano & Tomonaga, 2009](#), [2010](#)),
110 this study used the eye-tracking method to perform direct comparisons among humans, gorillas, and
111 orangutans with regard to their patterns of face and eye scanning. We employed a cross-species design,
112 presenting both conspecific and allospecific pictures. We initially presented whole-body pictures and
113 examined how humans and apes scanned faces versus bodies. We then presented facial pictures and
114 examined how they scanned each facial feature (e.g., eyes, nose, and mouth). Picture models included
115 both conspecific and allospecific individuals, males and females (adults), and familiar and unfamiliar
116 individuals. To examine the effect of species-specific facial morphologies on viewing patterns,
117 juvenile faces of gorillas and orangutans were prepared in addition to male adult and female adult
118 faces, and the viewing patterns for all types of faces were compared with one another.

Method

Participants

Five gorillas (*Gorilla gorilla*), 10 orangutans (*Pongo pygmaeus*), and 12 humans (five males and seven females; all European adults; 21–52 years, mean: 30.5 years) participated in this study. All apes were housed in semi-natural indoor and outdoor enclosures (total 2,564 m² for gorillas, 1,910 m² for orangutans) at the Wolfgang Köhler Primate Research Center (WKPRC), Leipzig Zoo, Germany. All apes received regular food, enrichment, and water *ad libitum*. They were not deprived of food or water. All apes and humans voluntarily participated in the study. Animal husbandry and research complied with the EAZA Minimum Standards for the Accommodation and Care of Animals in Zoos and Aquaria and the WAZA Ethical Guidelines for the Conduct of Research on Animals by Zoos and Aquariums, respectively. Informed consent was obtained from all human participants. Table 1 lists the sex, age, and rearing history of each ape. Most of the ape participants were reared by their biological mothers. All apes and humans had extensive and regular experience interacting with both apes and humans at the zoo and were thus highly familiar with both kinds of faces.

Table 1 about here

Figure 1 about here

Apparatus

All apes were tested in a chamber that separated the ape from the eye-tracking apparatus and the experimenter with transparent acrylic panels. The gaze movements of the apes were non-invasively recorded while their heads were unrestrained. We were able to implement this

unrestrained eye-tracking method by using two devices. First, we employed a table-mounted infrared eye tracker with wide-angle lenses (± 40 degrees in the semicircle above the camera; 60 Hz; Tobii X120, Tobii Technology AB, Stockholm, Sweden). This eye tracker recorded both eyes of participants (the average value was used to estimate the single gaze point) and allowed relatively large head movements by participants. Second, the eye tracker and the 17-inch LCD monitor (1280×1024 pixels) were mounted on a movable platform, and the distance between the platform and the participants was adjusted by the experimenter to the point at which the gaze could be most accurately recorded (approx. 60 cm). This adjustment was performed before each trial if necessary (but not during the trial).

To encourage the apes to sit still in front of the acrylic panel and face the eye tracker, the experimenter offered small pieces of fruit *ad libitum*. These were given to the apes before each trial if necessary (but not during the trial). However, eight of the 15 apes tested (Gorgo, Louna, Zola, Batak, Kila, Maia, Suaq, and Tanah; mostly juveniles) had difficulty in approaching the acrylic panel upon the request of the experimenter. We thus used a nozzle and tube attached to the acrylic panel, which produced regular drops of grape juice during the experiment; this motivated the apes to keep their heads in front of acrylic panel (Fig. 1a and b). The eye tracker was unable to track both eyes of one male adult gorilla (Gorgo) due to the wide distance between his eyes, and of one juvenile female gorilla (Louna) due to her strabismus. Therefore, we concealed the left eye of these apes from the eye tracker by attaching opaque tape to the panel and tracked only the right eye. Although these apes viewed stimuli only with their right eyes, they still had a full view of all stimuli with the typical 60-cm viewing distance. The preliminary analysis revealed no specific pattern of face and eye scanning in these two apes, so we did not exclude their data from the analysis. Although experiments were conducted for apes in the presence of the experimenter, the apes rarely attended to the experimenter during the presentation of stimuli. Also, the apes who received grape juice during the presentation of

167 pictures did not look at the nozzle but freely viewed the pictures.

168 Humans were tested using the same apparatus but in another room with no panel between
169 the eye tracker and participants. Our preliminary tests for accuracy revealed that the acrylic panels
170 (1.5–2 cm thick, with no scratches or dirt) had no influence on the eye-tracking data. Each
171 participant's gaze was recorded as a relative coordinate with respect to the monitor size (i.e., not as the
172 gaze angle). One degree of gaze angle corresponded to approximately 1 cm on the screen at a typical
173 60-cm viewing distance.

174 Calibration

175 An automated calibration process was employed. Several reference points were presented
176 sequentially at different locations on the screen to guide the participants to look at those points. The
177 system then automatically matched the raw-gaze data of participants with those reference points.
178 Humans were instructed to view the small dots that appeared on the screen, whereas a short video clip
179 or a piece of fruit was presented to attract the apes' gaze. Five reference points were used for humans,
180 whereas two were used for apes to reduce the time required for each calibration process.

181 After completing the calibration process, the calibration accuracy at five points on the
182 screen was inspected by attracting the participants' gaze to those points in the same way. The
183 calibration process was repeated if necessary (the calibration was usually repeated more frequently for
184 apes than for humans because of the limited calibration conditions mentioned above.). The accuracy
185 check was conducted before every session and occasionally during the session (between the trials). In
186 a preliminary session, we estimated the error value (the average distance between recorded and
187 intended gaze positions) with two gorillas, six orangutans, six chimpanzees, and six humans. We
188 found that the error was small and comparable among the species (average \pm s.e.m. 0.61 ± 0.06 , $0.72 \pm$
189 0.04 , and 0.52 ± 0.05 degrees at the typical 60-cm viewing distance, respectively).

Figure 2 about here

Stimuli

Figures 2a and 2b present examples of the stimuli. Color still pictures were prepared for this study, including 16 whole-body and 56 facial pictures. Gorillas and orangutans viewed eight human and eight conspecific whole-body pictures (total of 16 whole-body pictures) and eight human and 24 conspecific facial pictures (32 facial pictures); each ape viewed a total of 48 pictures. Humans viewed pictures of both apes and humans (16 whole-body and 56 facial pictures). Half of the pictures were of familiar apes/humans that interacted with the participants on a daily basis, and the other half included unfamiliar apes/humans that had been never exposed to the participants. Pictures were taken when the models were in calm, relaxed states. Their eyes were open, and their mouths were closed. The facial pictures of gorillas and orangutans consisted of three types of faces (eight stimuli for each type): adult male face, adult female face, and juvenile face. Whole-body pictures were converted to 1280×1024 pixels (37×30 degrees at a typical 60-cm viewing distance). Facial pictures were converted to 768×1024 pixels (22×30 degrees at a typical 60-cm viewing distance) with a gray frame around the background (total 1280×1024 pixels). The internal parts of faces (eyes, nose, and mouth) were thus approximately 10–15 degrees in size at a typical 60-cm viewing distance.

Procedure

In each trial, a picture was presented after participants fixated on a red mark that appeared at the center of the screen. Participants then scanned the picture freely (without any training or instruction). They never kept gazing at the point where the initial red mark appeared, and sequential scanning of the picture was almost always observed. Each stimulus was presented for 3 sec. No

specific instructions were given to humans except to view the pictures freely.

The presentation order of the pictures was randomized for each participant. The entire session was conducted over 2 days for humans (36 pictures each day), but the session was divided across 12 days for apes (four pictures each day). The purpose of dividing the entire session for apes was to reduce the time required for daily sessions and to maintain the apes' interest in the pictures. Daily sessions lasted 10–15 min for apes and 15–20 min for humans. Trials in which participants viewed a picture for less than 1 sec were repeated after the whole session, and the original trials were replaced by the new trials; otherwise, those trials were eliminated from the analysis. As a result, we excluded 6.2% and 8.7% of the trials of the whole-body pictures and 10.0% and 7.1% of the trials of the facial pictures for gorillas and orangutans, respectively (no trials were eliminated for humans). The exclusion of trials (especially those of facial pictures) was largely attributable to two male human-raised individuals, Gorgo (gorilla) and Bimbo (orangutan). These individuals sometimes averted their heads from facial pictures. The reason for this behavior is uncertain; it could have been active gaze avoidance in response to the social stimuli or simply a lack of interest in the pictures.

Data analysis

Fixation definition

Fixation was defined as a stationary gaze within a radius of 50 pixels for at least 75 ms (more than five measurement samples). Otherwise, the recorded sample was defined as part of a saccade. Records during the first 200 ms were eliminated from the analysis, thereby eliminating fixations that might have begun before the onset of stimuli.

Area of interest (AOI)

Each stimulus was divided into areas of interest (AOI) for quantitative comparison. Each whole-body picture was divided into background, face, and body. Each AOI was defined to be 20

pixels larger than the precise outline of the features to compensate for error in gaze estimation. The AOIs were generated in the following order: background, body, and face. If two or more AOIs were duplicated, gaze samples were added to the last AOI. Each facial picture was divided into background, eyes, nose, mouth, and periphery (ears, cheeks, chin, forehead, hair; see Fig. 2c).

Proportion of viewing time

The proportion of viewing time for each AOI was calculated with respect to the viewing time for the entire scene. Out-of-scene fixations were excluded from the analyses (less than 5% of all fixations in all species when participants were presented with whole-body pictures; 6.1%, 6.0%, and 1.9% of all fixations in gorillas, orangutans, and humans, respectively, when participants were presented with facial pictures). To compensate for the differences in area size between AOIs, viewing time was normalized for area size by subtracting the proportion of viewing time from the proportion of area size. The chance level was thus set at zero.

Probability of fixation across fixation order

To examine the time course of face viewing in whole-body pictures and of eye viewing in facial pictures, we calculated the probability of fixation on faces/eyes across fixation order (first through fifth fixations). The results were calculated as the proportion of fixations with respect to the total number of sampled fixations for each fixation order.

Results

Figure 3 about here

Figure 3 shows examples of scanning paths in each species. First, it is important to note that none of the ape participants exhibited a fear response to facial pictures, even though some apes (especially juveniles) had never been exposed to facial pictures before this study.

Figure 4 about here

Whole-body pictures

Proportion of viewing time. Figure 4a shows the proportion of viewing time for each AOI when participants were presented with whole-body pictures. We conducted a repeated-measures ANOVA with species, AOI, stimulus species, and familiarity as factors. We found a significant main effect of AOI ($F(2, 48) = 446.59, P < 0.001, \eta^2 = 0.94$). All species viewed each AOI in a different way [humans ($F(2, 22) = 382.82, P < 0.001, \eta^2 = 0.97$), gorillas ($F(2, 8) = 50.41, P < 0.001, \eta^2 = 0.92$), orangutans ($F(2, 18) = 178.01, P < 0.001, \eta^2 = 0.95$)]. *Post hoc* tests (Bonferroni correction) revealed that they viewed faces most intensely, followed by bodies and then backgrounds ($P < 0.001$). We found a significant interaction between species and AOI ($F(4, 48) = 18.64, P < 0.001, \eta^2 = 0.60$). We found significant species differences in viewing patterns for faces ($F(2, 24) = 28.69, P < 0.001, \eta^2 = 0.70$), bodies ($F(2, 24) = 5.15, P = 0.014, \eta^2 = 0.30$), and backgrounds ($F(2, 24) = 24.50, P < 0.001, \eta^2 = 0.67$). This is explained by the fact that humans viewed faces more intensely than did apes ($P < 0.001$). Although we were not able to analyze the effects of age and sex because of the small numbers of juveniles and males, excluding juveniles or males from the analysis did not change the

overall pattern of results [juveniles (main AOI: $F(2, 40) = 396.84, P < 0.001, \eta^2 = 0.95$; species \times AOI: $F(4, 40) = 20.20, P < 0.001, \eta^2 = 0.66$), males (main AOI: $F(2, 44) = 391.20, P < 0.001, \eta^2 = 0.94$; species \times AOI: $F(4, 44) = 17.34, P < 0.001, \eta^2 = 0.61$)]. To examine whether participants showed differential patterns over the course of an entire session, we separated the whole session into two blocks and included that factor in the ANOVA. However, we did not find any effect of session block (block \times AOI: $F(2, 23) = 0.76, P = 0.47, \eta^2 = 0.06$; block \times AOI \times species: $F(4, 48) = 1.19, P = 0.32, \eta^2 = 0.09$).

We found a significant interaction between stimulus species and AOI ($F(2, 48) = 4.72, P = 0.013, \eta^2 = 0.16$). The effect of stimulus species was significant in viewing patterns for faces ($F(1, 24) = 10.56, P = 0.003, \eta^2 = 0.30$) and bodies ($F(1, 24) = 6.09, P = 0.021, \eta^2 = 0.20$). This is explained by the fact that all species viewed conspecific faces longer than allospecific faces. Additionally, we found a significant interaction between familiarity, species, and AOI ($F(4, 48) = 2.93, P = 0.030, \eta^2 = 0.19$). The interaction between familiarity and AOI was significant only in humans ($F(2, 22) = 4.34, P = 0.026, \eta^2 = 0.28$). Humans viewed faces of familiar individuals somewhat longer than those of unfamiliar individuals (0.65 vs. 0.60).

Probability of fixation on faces across fixation order. Figure 4b shows the time course of face-viewing patterns for whole-body pictures. A repeated-measures ANOVA with species and fixation order as factors revealed a main effect of fixation order ($F(4, 96) = 18.23, P < 0.001, \eta^2 = 0.43$). This is explained by the fact that they fixated on faces at an earlier rather than a later time. We also found a significant interaction between species and fixation order ($F(8, 96) = 4.22, P < 0.001, \eta^2 = 0.26$). This is explained by the fact that humans fixated on faces at an earlier time than did apes.

Figure 5 about here

Facial pictures

Proportion of viewing time. Figure 5a shows the proportion of viewing time for each AOI when participants were presented with facial pictures. We conducted a repeated-measures ANOVA with species, AOI, stimulus species, and familiarity as factors. Only adult faces of the three species were used in this analysis. We found a significant main effect of AOI ($F(3, 69) = 238.75, P < 0.001, \eta^2 = 0.91$). All species viewed each AOI in a different way; humans ($F(3, 33) = 146.57, P < 0.001, \eta^2 = 0.93$), gorillas ($F(3, 12) = 100.94, P < 0.001, \eta^2 = 0.96$), and orangutans ($F(3, 24) = 65.90, P < 0.001, \eta^2 = 0.89$). *Post hoc* tests (Bonferroni correction) revealed that they viewed eyes most intensely, followed by nose/mouth and periphery ($P < 0.001$). We found a significant interaction between species and AOI ($F(3, 69) = 3.14, P = 0.009, \eta^2 = 0.21$). We found significant species differences in viewing patterns for the eyes ($F(2, 23) = 4.00, P = 0.032, \eta^2 = 0.25$) and the periphery ($F(2, 23) = 6.27, P = 0.007, \eta^2 = 0.35$). This is explained by the fact that humans viewed the eyes more intensely and viewed the periphery less intensely than did orangutans. Although we were not able to analyze the effects of age and sex because of the small numbers of juveniles and males, excluding juveniles or males from the analysis did not change the overall pattern of results [juveniles (main AOI: $F(3, 57) = 174.85, P < 0.001, \eta^2 = 0.90$; species \times AOI: $F(6, 57) = 3.20, P = 0.009, \eta^2 = 0.25$), males (main AOI: $F(3, 66) = 218.51, P < 0.001, \eta^2 = 0.90$; species \times AOI: $F(6, 66) = 3.13, P = 0.009, \eta^2 = 0.22$)]. To examine whether the participants showed differential patterns over the course of entire session, we separated the whole session into two blocks and included that factor in the ANOVA. However, we did not find any effect of session block (block \times AOI: $F(3, 22) = 1.87, P = 0.16, \eta^2 = 0.20$; block \times AOI \times species: $F(3, 22) = 1.49, P = 0.20, \eta^2 = 0.16$).

We found a significant interaction among stimulus species, AOI, and species ($F(6, 69) = 6.62, P < 0.001, \eta^2 = 0.36$). This is explained by the fact that, whereas gorillas and orangutans viewed the conspecific eyes more intensely than the allospecific eyes, humans exhibited an opposite tendency,

viewing the allospecific eyes more intensely than the conspecific eyes. The effect of familiarity was not significant ($P > 0.05$).

Probability of fixation on eyes across fixation order. Figure 5b shows the time course of eye viewing for facial pictures. A repeated-measures ANOVA using species and fixation order as factors revealed a main effect of fixation order ($F(4, 96) = 35.03, P < 0.001, \eta^2 = 0.59$), which is explained by the fact that they fixated on eyes at an earlier rather than a later time. The interaction between species and fixation order was not significant ($F(8, 96) = 1.69, P = 0.11, \eta^2 = 0.12$).

Viewing patterns for the eyeball region. Differences between apes and humans in eye viewing were not as clear as those previously found between chimpanzees and humans (Kano & Tomonaga, 2009, 2010). However, this possible inconsistency may be attributable to the definition of the eye AOI. That is, the visual inspection of each scanning path (Fig. 3) suggested that humans more directly fixated on eyes than did apes (i.e., fixation on eyeballs rather than on the region around the eyes). Therefore, we also delineated an eyeball AOI, which was inside the eye AOI and included only the eyeball, and compared the viewing times for eyeballs between the species. We found significant species differences in viewing patterns for eyeballs (Fig. 5a; $F(2, 23) = 8.09, P = 0.002, \eta^2 = 0.41$). This is explained by the fact that humans viewed eyeballs more intensely than did apes ($P < 0.05$). We also analyzed the probability of fixation on eyeballs across fixation order. We found a significant interaction between species and fixation order (Fig. 5b; $F(8, 96) = 2.28, P = 0.028, \eta^2 = 0.16$). This is explained by the fact that, whereas humans showed an increasing tendency for viewing eyeball regions, apes showed a decreasing tendency in this regard.

Figure 6 about here

Probability of saccade destination. Visual inspection of the scanning data (Fig. 3) suggested

that humans viewed left and right eyes alternately, whereas apes did not. To clarify this issue, we quantified the typical scanning path of each species when participants were presented with facial pictures. We divided the eye AOI into right and left eye AOIs (i.e., right eye, left eye, nose, mouth, and periphery) and calculated the probability of saccade destination. Each saccade (the path that joins two consecutive fixations) was classified based on the two AOIs in which the saccade started and ended; thus, each saccade was classified into one of 10 possible combinations of beginning and ending AOIs. Saccades for out-of-face start or end points were excluded from analyses. The results were calculated as the proportion of the total number of sampled saccades. Data for conspecific and allospecific faces were combined. As suggested, humans shifted their gaze between the left and right eyes more frequently than did gorillas ($t(15) = 2.26, P = 0.039$) and orangutans ($t(20) = 4.82, P < 0.001$) (Fig. 6).

Figure 7 about here

The effect of face type. Figure 7 shows the proportion of viewing time for each AOI as a function of face type in gorillas and orangutans. A repeated-measures ANOVA with face type and AOI did not reveal an effect of face type in gorillas ($F(6, 24) = 1.97, P = 0.10, \eta^2 = 0.33$), whereas it did in orangutans ($F(6, 54) = 6.09, P < 0.001, \eta^2 = 0.40$). *Post hoc* tests revealed that orangutans viewed the juvenile eyes more intensely than the male or female eyes ($P < 0.05$). We also found that they viewed the periphery of male faces (i.e., cheek flange) more intensely than that of female or infant faces ($P < 0.05$). Human participants did not demonstrate this pattern (gorilla faces: $F(6, 66) = 0.08, P = 0.99, \eta^2 = 0.008$; orangutan faces: $F(6, 66) = 1.53, P = 0.18, \eta^2 = 0.12$).

375 Discussion

376 First, it is important to note that humans and great apes showed striking similarities in their
377 patterns of face and eye scanning. All species viewed faces for longer durations and at earlier times
378 than they viewed bodies and backgrounds when presented with whole-body pictures. Additionally, all
379 species viewed inner features of faces, especially eyes, for longer durations and at earlier times than
380 they viewed the periphery when presented with facial pictures. The same patterns were observed for
381 chimpanzees in previous studies (Kano & Tomonaga, 2009, 2010) (Table S1 and S2). Selective
382 attention to faces as opposed to complex naturalistic backgrounds clearly eliminated the possibility
383 that viewing patterns were simply dependent on visual saliency (e.g., color, shape, contrast). In
384 support of this notion, a previous study demonstrated that the chimpanzees' viewing patterns for faces
385 were not explained by the saliency model of Itti and Koch (2001) (Kano & Tomonaga, 2011).

386 None of the ape species avoided viewing eyes in this study. Rather, they viewed eyes
387 predominantly, regardless of the type of face (conspecific/allospecific, male/female) presented. A
388 possible inconsistency between these data and previous findings that gorillas and orangutans
389 frequently avoided viewing faces/eyes during natural interactions may exist (Kaplan & Rogers, 2002;
390 Shaller, 1963). However, this apparent inconsistency may be explained by the fact that this study
391 presented apes with faces in the absence of interactive contexts. Previous studies have also shown that
392 eye contact (or its avoidance) is highly dependent on the interactive contexts in which such eye contact
393 occurs. Thus, this study's finding that all great ape species are similar in their eye-viewing tendencies
394 may indicate that the ape species do not differ in their default motivation for viewing eyes. However, it
395 should be noted that two adult males in this study frequently averted their heads when presented with
396 facial pictures (see Methods). It is possible that notable sex, age, rank, or individual differences
397 characterize their eye-viewing tendencies (cf. Yamagiwa, 1992)

398 Although no significant differences were found between gorillas and orangutans, humans

differed from apes in their patterns of face and eye scanning. Humans viewed faces for longer durations than did apes when presented with whole-body pictures. Additionally, humans viewed eyes, especially eyeballs, for longer durations than did apes when presented with facial pictures. Furthermore, humans often alternated their gaze between the left and right eyes, whereas apes rarely showed this gaze movement (Fig. 3, 6). By alternating their gaze between the left and right eyes and shifting their gaze occasionally to the mouth, humans exhibited triangular scanning paths in response to faces. As the upper component of this triangular shape was absent in apes, their scanning paths appear to be more linear than those of humans. A similar species difference was observed between chimpanzees and humans in previous studies (Kano & Tomonaga, 2009, 2010). Therefore, active, prolonged eye viewing is a unique characteristic of humans as compared with great apes.

One possible interpretation for this species difference is that eyes may have evolved additional communicative functions not found in ape species. Compared with those of other primates, human eyes have a notable dark–white contrast between the iris and sclera (Kobayashi & Kohshima, 2001), which enhances directional cues. Moreover, fine motor control of the muscles around the eyes (Ekman & Friesen, 1978) enables humans to communicate a variety of emotional cues. As these cues are subtle, the active and prolonged eye-viewing patterns of humans may benefit them by facilitating the efficient retrieval of such communicative information. That is, humans may employ eyes for bidirectional communication between senders and receivers. Tomasello et al. (2007) hypothesized that the importance of close-range joint attentional and communicative interactions may have shaped the unique eye morphology of humans. It is also possible that the behavioral tendency to actively collect eye information has evolved along with eye morphology.

Humans and apes viewed conspecific faces for longer durations than they viewed allospecific faces when presented with whole-body pictures. This may indicate their greater interest in conspecific than in allospecific individuals. However, their patterns of face and eye scanning were, in general,

highly consistent across stimuli despite the fact that various types of faces were presented. This was also the case for chimpanzees and humans in the previous study. Thus, their viewing patterns may reflect general responses to face-like configurations rather than specific responses to particular facial information. This result may be explained by the fact that study participants were not required to collect particular information from the faces (free viewing). Thus, we would expect that their viewing patterns would be more variable if they viewed faces under an experimental situation in which they were required to sort faces based on certain facial information or in an interactive situation in which they reacted emotionally to faces. Further studies are necessary to test these possibilities.

Although orangutans' viewing patterns were generally consistent across stimuli, they were clearly affected by species-specific facial features. They viewed the periphery of adult male faces (flange on the sides) for longer durations than those of adult female or juvenile faces. Furthermore, they viewed the eyes of juvenile faces (surrounded by pale coloring) for longer durations than those of adult faces. These results suggest that these conspicuous facial features are indeed visually appealing. Human participants who viewed the same orangutan faces did not follow the same patterns, perhaps because their viewing patterns are more standardized than are those of orangutans.

The color contrast in human eyes did not attract viewers' attention to the eyes. That is, apes did not view human eyes for longer durations than they viewed conspecific eyes, and humans showed prolonged eye viewing for both conspecific and allospecific faces. Thus, unlike the male flange and the skin color of juvenile orangutans, the color contrast in human eyes is not particularly conspicuous. Humans showed prolonged eye viewing independently of the effect of eye color. Thus, this pattern is internally driven rather than stimulus dependent. We speculate that the color contrast of human eyes conveys information most efficiently when the eyes receive such active attention.

Further comparative studies on face and eye scanning may reveal similarities and differences between apes and other primate species. For example, this study demonstrated a remarkably strong

tendency for apes and humans to view the internal features of faces, and successive fixations on internal features were frequent (i.e., successive on-feature fixations rather than repetition of on- and off-feature fixations; see Fig. 3). It remains unclear how apes and monkeys differ in this regard when tested using the same experimental settings. Further studies using the eye-tracking method are also necessary to perform direct comparisons between humans and great apes in terms of gaze-following tendencies. Previous studies have shown that humans, chimpanzees, gorillas, and orangutans show differential gaze-following tendencies in response to a human experimenter's gaze (Okamoto-Barth, et al., 2007; Tomasello, et al., 2007). It remains unclear how these species differ from one another when tested using a cross-species design (presenting both conspecific and allospecific faces).

The aim of this study was to reveal how humans and great apes are similar and different in their pattern of face and eye scanning. Following a previous study comparing chimpanzees with humans, we directly compared the viewing patterns of gorillas, orangutans, and humans in response to whole-body and facial (full-face) pictures of conspecifics and allospecifics. The general conclusion is that all species are strikingly similar in patterns of face and eye scanning. However, we also identified unique eye-viewing patterns among humans. Unlike the species-specific facial features of orangutans (e.g., male flange), the black–white contrast in human eyes, a unique trait of humans, did not attract viewers' attention. Humans showed prolonged eye viewing independently of the eye color of presented faces. Thus, active eye viewing may be a unique characteristic of humans as compared with great apes, and given that humans have unique eye morphology, facial communication among humans may be specialized for the eyes.

Acknowledgements

This research was financially supported by the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan Grants-in-Aid

471 for Scientific Research (nos. 16002001, 19300091, 20002001, 212299), the JSPS/MEXT global COE
472 programs (D07 and A06), and the JSPS institutional program for young researchers overseas. We
473 thank C. Rahn, F. Stock, H. Knoke, J. Corbit, R. Pieszek, and the other members of WKPRC for their
474 help with experiments. We also thank the Leipzig Zoo, Germany, and the Centre for Human Evolution
475 Modelling Research at the Primate Research Institute, Japan, for the daily care of the apes. Pictures of
476 apes are courtesy of A. Bania, A. Girndt, A. Sato, J. Call, J. Kalbitz, N. Kuze, S. Ross, and T. Stoinski.
477

478 References

- 479 Ankel-Simons, F. (2000). *Primate anatomy: An introduction*. San Diego: Academic Press.
- 480 Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and
481 around barriers. *Journal of Comparative Psychology*, 119(2), 145–154.
- 482 Chevalier-Skolnikoff, S. (1973). Facial expression of emotion in non-human primates. In P. Ekman
483 (Ed.), *Darwin and facial expression: A century of research in review* (pp. 11–89). New York:
484 Academic Press.
- 485 Darwin, C. (1999). *The expression of the emotions in man and animals* (P. Ekman, Ed). London:
486 Fontana. (Original work published 1872).
- 487 de Waal, F. B. M. (1990). *Peacemaking among primates*. Cambridge: Harvard University Press.
- 488 Ekman, P., & Friesen, W. V. (1978). *The facial action coding system (FACS): A technique for the*
489 *measurement of facial action*. Palo Alto: Consulting Psychologists Press.
- 490 Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze.
491 *Neuroscience and Biobehavioral Reviews*, 24(6), 581–604.
- 492 Gomez, J. C. (1996). Ostensive behavior in great apes: The role of eye contact. In A. E. Russon, K. A.
493 Bard, & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (pp. 131–151).
494 New York: Cambridge University Press.
- 495 Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge: Belknap Press.
- 496 Goodenough, J., McGuire, B., & Jakob, E. (1993). *Perspectives on animal behavior*. New York:
497 Wiley.
- 498 Itakura, S. (2004). Gaze-following and joint visual attention in non-human animals. *Japanese*
499 *Psychological Research*, 46(3), 216–226.
- 500 Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews*
501 *Neuroscience*, 2(3), 194–203.

- 502 Kano, F., & Tomonaga, M. (2009). How chimpanzees look at pictures: A comparative eye-tracking
503 study. *Proceedings of the Royal Society B: Biological Sciences*, 276(1664), 1949–1955.
- 504 Kano, F., & Tomonaga, M. (2010). Face scanning in chimpanzees and humans: Continuity and
505 discontinuity. *Animal Behaviour*, 79(1), 227–235.
- 506 Kano, F., & Tomonaga, M. (2011). Perceptual mechanism underlying gaze guidance in chimpanzees
507 and humans. *Animal Cognition*, 14(3), 377–386.
- 508 Kano, T. (1980). Social behavior of wild pygmy chimpanzees (*Pan paniscus*) of Wamba: A
509 preliminary report. *Journal of Human Evolution*, 9(4), 243–260.
- 510 Kaplan, G., & Rogers, L. J. (2002). Patterns of gazing in orangutans (*Pongo pygmaeus*). *International*
511 *Journal of Primatology*, 23(3), 501–526.
- 512 Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, 100(1),
513 78–100.
- 514 Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive
515 meaning: Comparative studies on external morphology of the primate eye. *Journal of Human*
516 *Evolution*, 40(5), 419–435.
- 517 Kuze, N., Malim, T. P., & Kohshima, S. (2005). Developmental changes in the facial morphology of
518 the Borneo Orangutan (*Pongo pygmaeus*): Possible signals in visual communication.
519 *American Journal of Primatology*, 65(4), 353–376.
- 520 Okamoto-Barth, S., Call, J., & Tomasello, M. (2007). Great apes' understanding of other individuals'
521 line of sight. *Psychological Science*, 18(5), 462–468.
- 522 Redican, W. K. (1975). Facial expressions in non-human primates. In L. A. Rosenblum (Ed.), *Primate*
523 *Behavior* (Vol. 4, pp. 103–194): Academic Press.
- 524 Shaller, G. B. (1963). *The mountain gorilla: Ecology and behavior*. Chicago: Chicago University
525 Press.

- 526 Thomsen, C. E. (1974). Eye contact by non-human primates toward a human observer. *Animal*
527 *Behaviour*, 22, 144–149.
- 528 Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze
529 following of great apes and human infants: The cooperative eye hypothesis. *Journal of*
530 *Human Evolution*, 52(3), 314–320.
- 531 Tomonaga, M. (2010). Chimpanzee eyes have it? Social cognition on the basis of gaze and attention
532 from the comparative–cognitive–developmental perspective. In E. Lornsdorf, S. Ross & T.
533 Matsuzawa (Eds.), *The mind of the chimpanzee: Ecological and empirical perspectives*.
534 Chicago: University of Chicago Press.
- 535 van Hooff, J. A. R. A. M. (1967). The facial displays of the catanrhine monkeys and apes. In D. Morris
536 (Ed.), *Primate ethology* (pp. 7–68). Piscataway, NJ: Aldine Transaction.
- 537 Yamagiwa, J. (1992). Functionial analysis of social staring behavior in an all-male group of mountain
538 gorillas. *Primates*, 33(4), 523–544.
- 539
- 540

Table 1. Age, sex, and rearing history of each ape.

	Sex	Age (years)	Rearing History
Gorilla			
Gorgo	M	29	Nursery
Kibara	F	6	Mother
Louna	F	4	Mother
Viringika	F	15	Mother
Zola	F	2	Mother
Orangutan			
Batak	M	1 (+ 1 month)	Mother
Bimbo	M	30	Nursery
Dukana	F	21	Mother
Kila	F	10	Mother
Maia	F	3	Mother
Padana	F	13	Mother
Pini	F	22	Mother
Raja	F	7	Mother
Suaq	M	1 (+ 6 months)	Mother
Tanah	F	1 (+ 5 months)	Mother

541

542

Figure captions

Figure 1. Apes and apparatus. (a) Eye-tracking apparatus. An eye tracker and a monitor are mounted on the movable platform (front). The experimenter controlled the apparatus using the computer mounted on the other platform (behind). See also:

<http://www.youtube.com/watch?v=zHx2KwQEHq0>.

(b) A juvenile gorilla drinking grape juice *via* a tube attached to the transparent acrylic panel. (c) An adult female orangutan on the apparatus.

Figure 2. Examples of stimuli: (a) whole-body pictures; (b) facial pictures; (c) areas of interest (AOI) for facial pictures.

Figure 3. Examples of scanning paths for facial pictures. The scanning paths of five participants from each species were superimposed on the facial pictures. The pictures were dimmed for clarity.

Figure 4. Scanning pattern for whole-body pictures in each species. (a) Proportion of viewing time for each AOI with respect to the total scene-viewing time. (b) Probability of fixation for each AOI across fixation order. All data are normalized for area sizes. Error bars represent the upper and lower bounds of the 95% confidence intervals.

Figure 5. Scanning pattern for facial pictures in each species. (a) Proportion of viewing time for each AOI with respect to total face-viewing time. (b) Probability of fixation for each AOI across fixation order. All data are normalized for area sizes. Error bars represent the upper and lower bounds of the 95% confidence intervals.

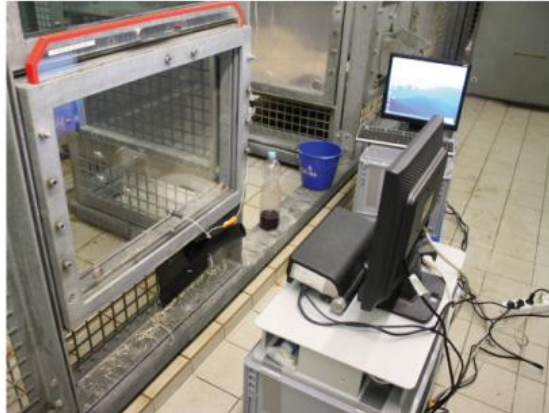
Figure 6. Typical scanning paths for facial pictures. Probabilities of saccade destination are presented in percentages and scaled to size. If a human participant is currently looking at the right eye of a human face (top center), he/she would re-fixate the right eye with a probability of 2.6, the left eye with a probability of 15.6, the nose with a probability of 7.7, etc. (see text for details).

Figure 7. The proportion of total face-viewing time spent viewing each AOI as a function of face type in gorillas and orangutans. All data were normalized for area sizes. Error bars represent the upper and lower bounds of the 95% confidence intervals. Asterisk indicates $P < 0.05$ in *post hoc* tests.

The English in this document has been checked by at least two professional editors, both native speakers of English. For a certificate, please see:

<http://www.textcheck.com/certificate/HCZkUh>

(a)



(b)



(c)



Figure 1

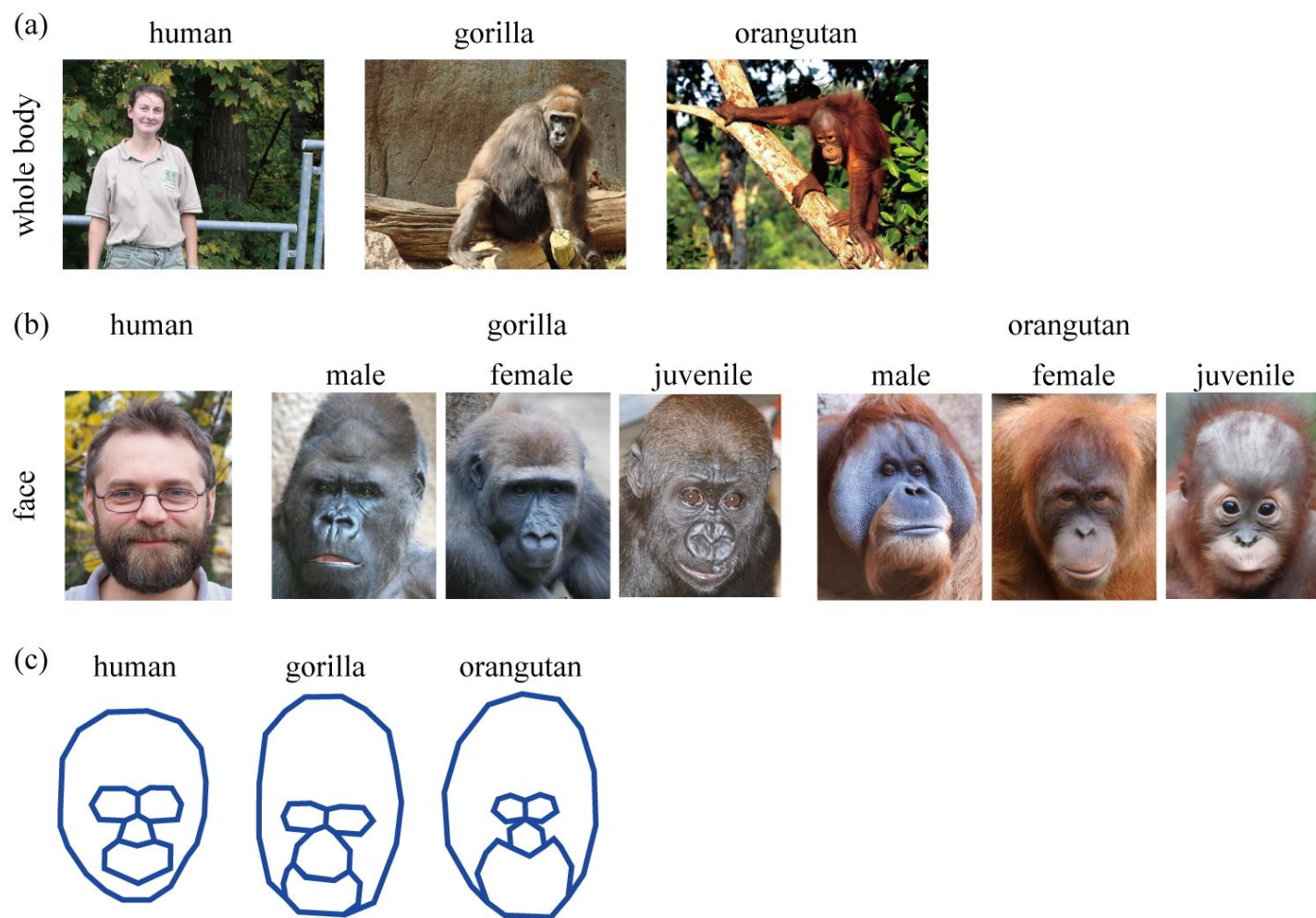
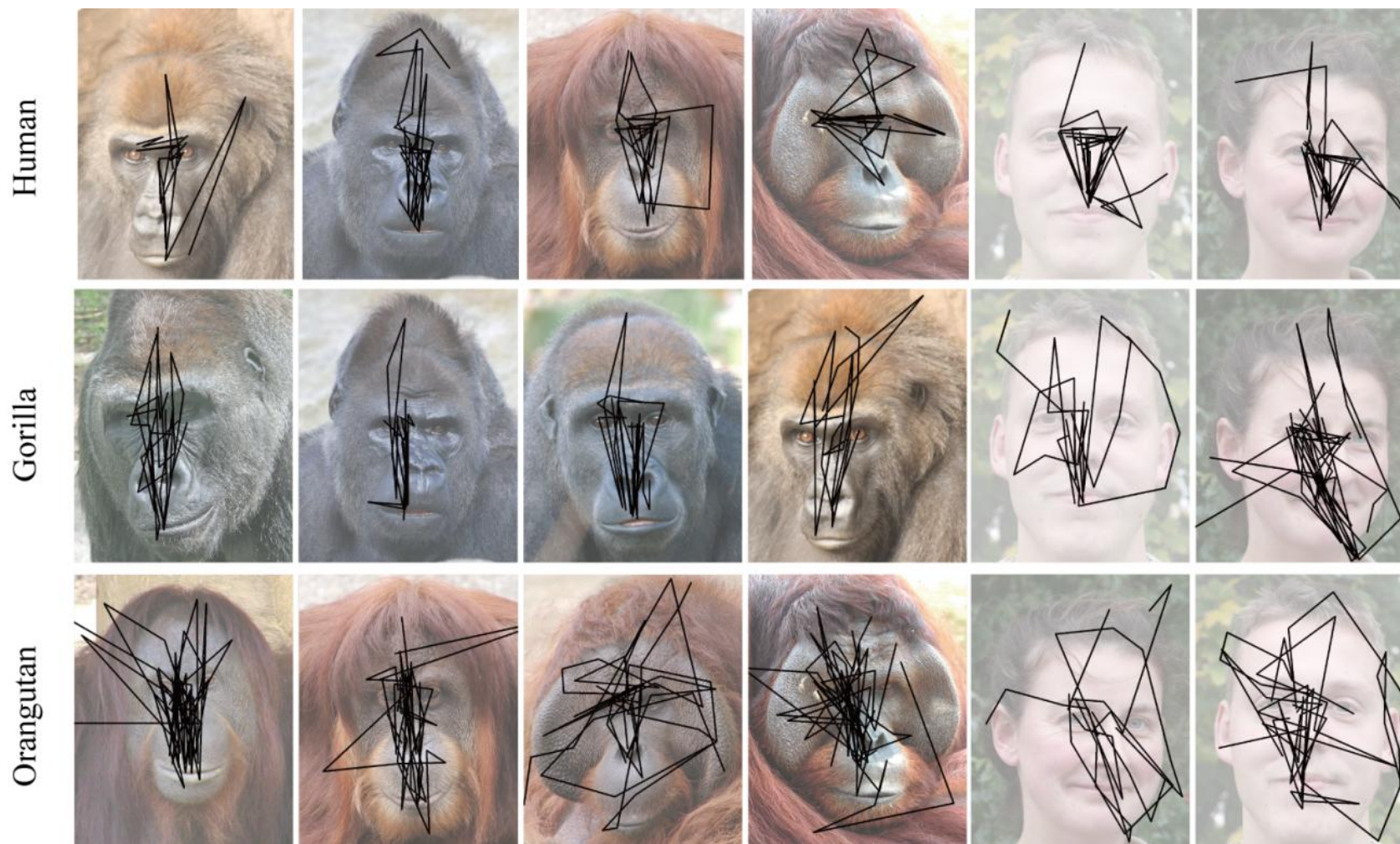


Figure 2



592

593

Figure 3

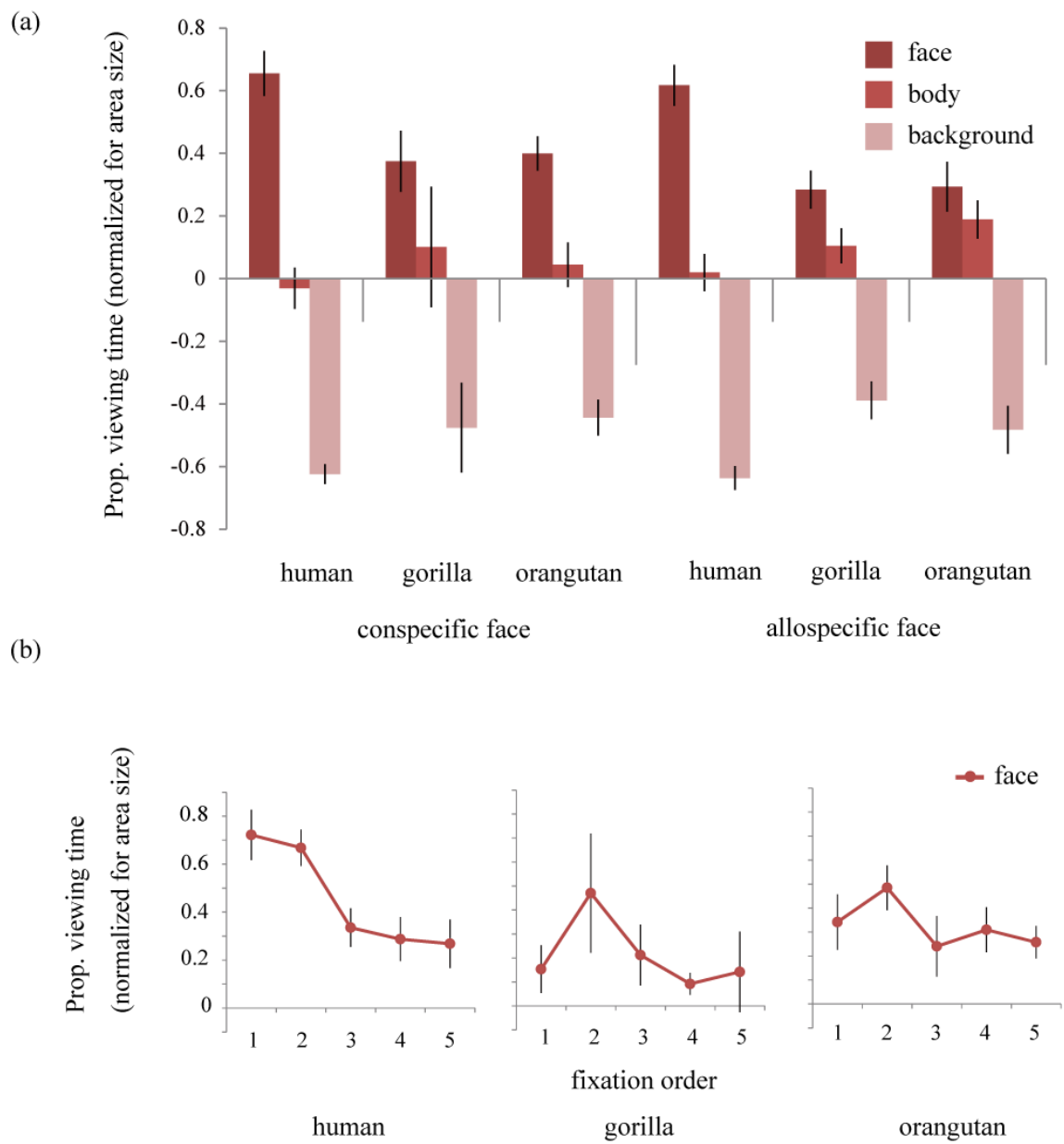


Figure 4

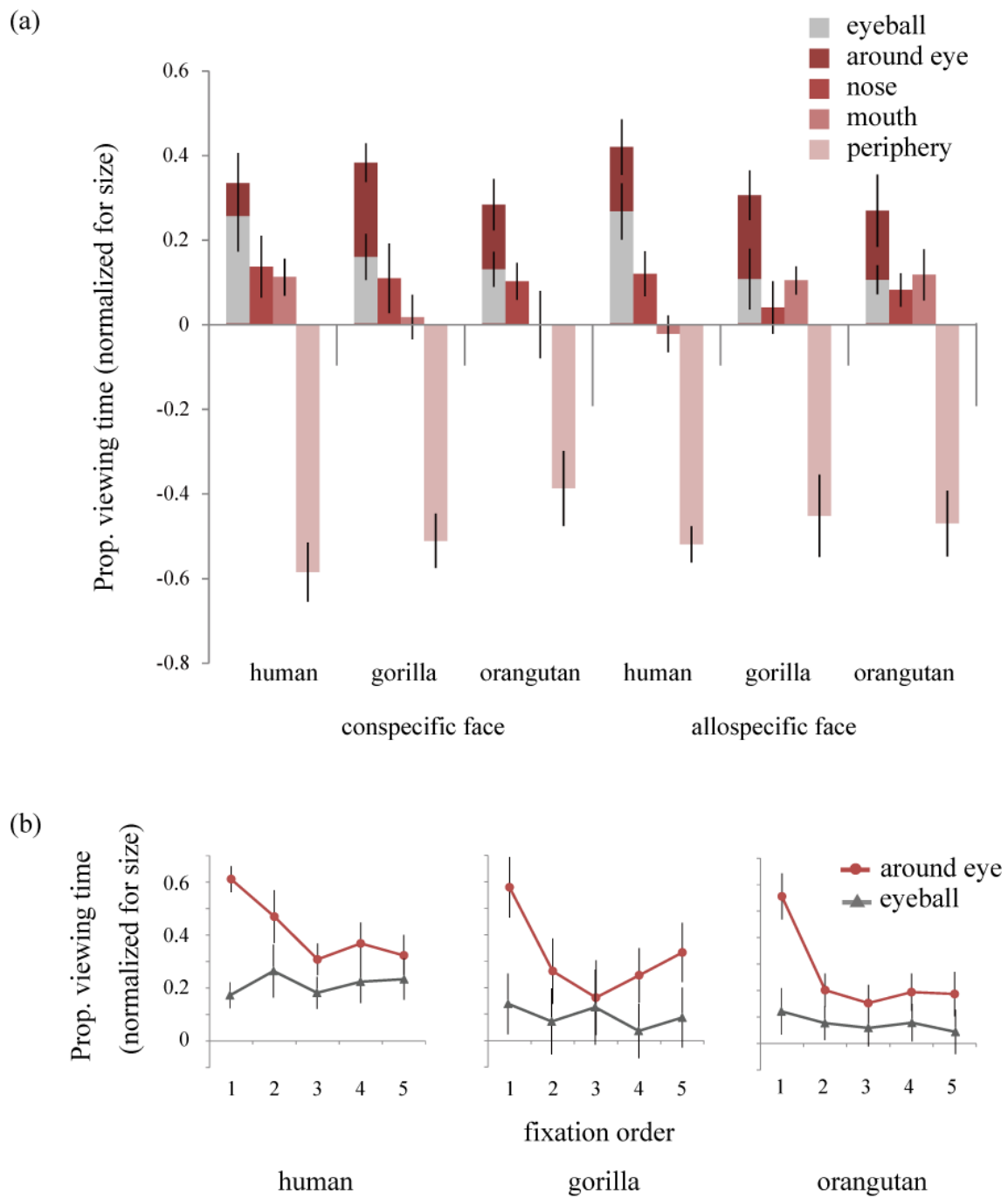


Figure 5

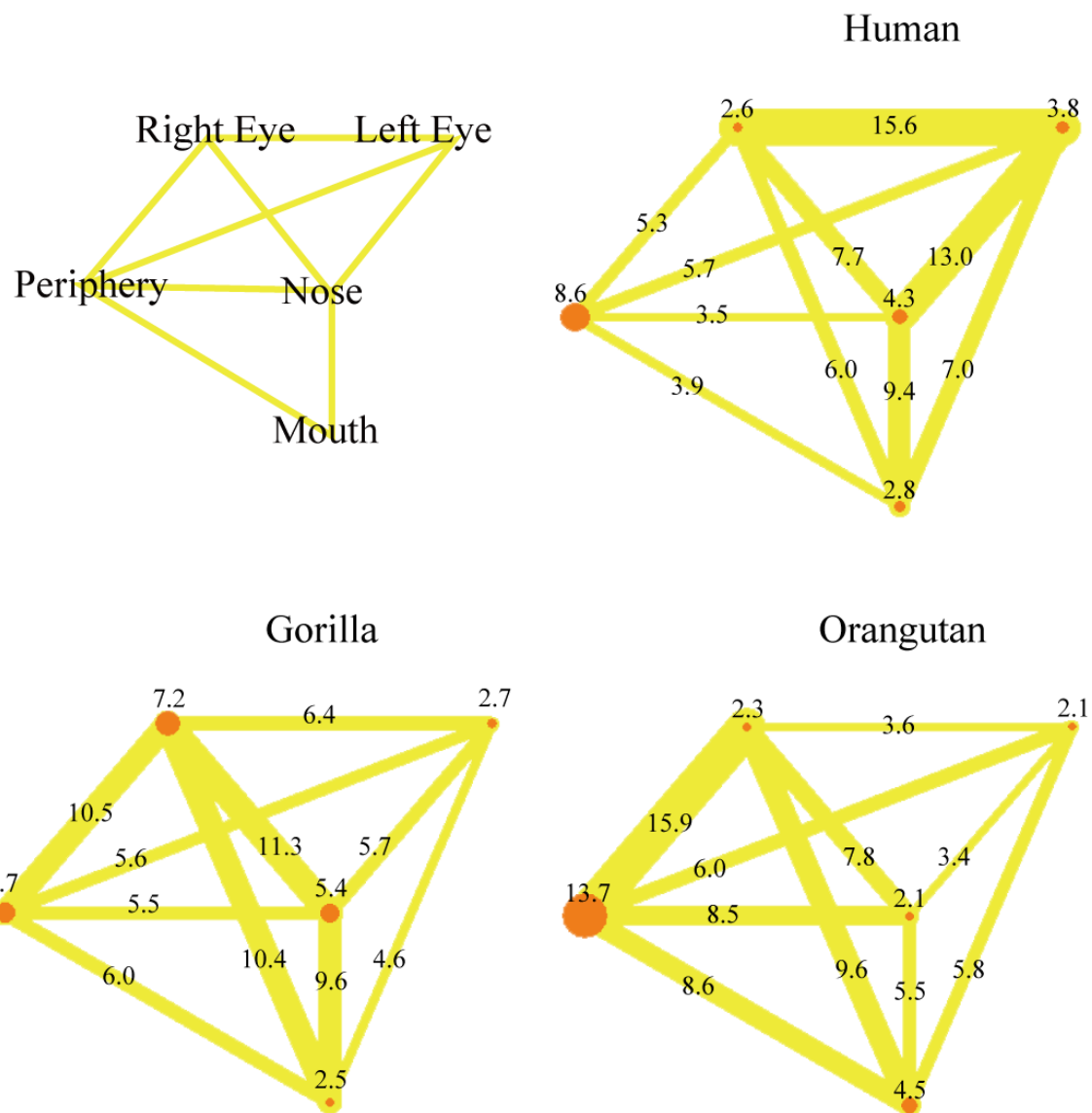
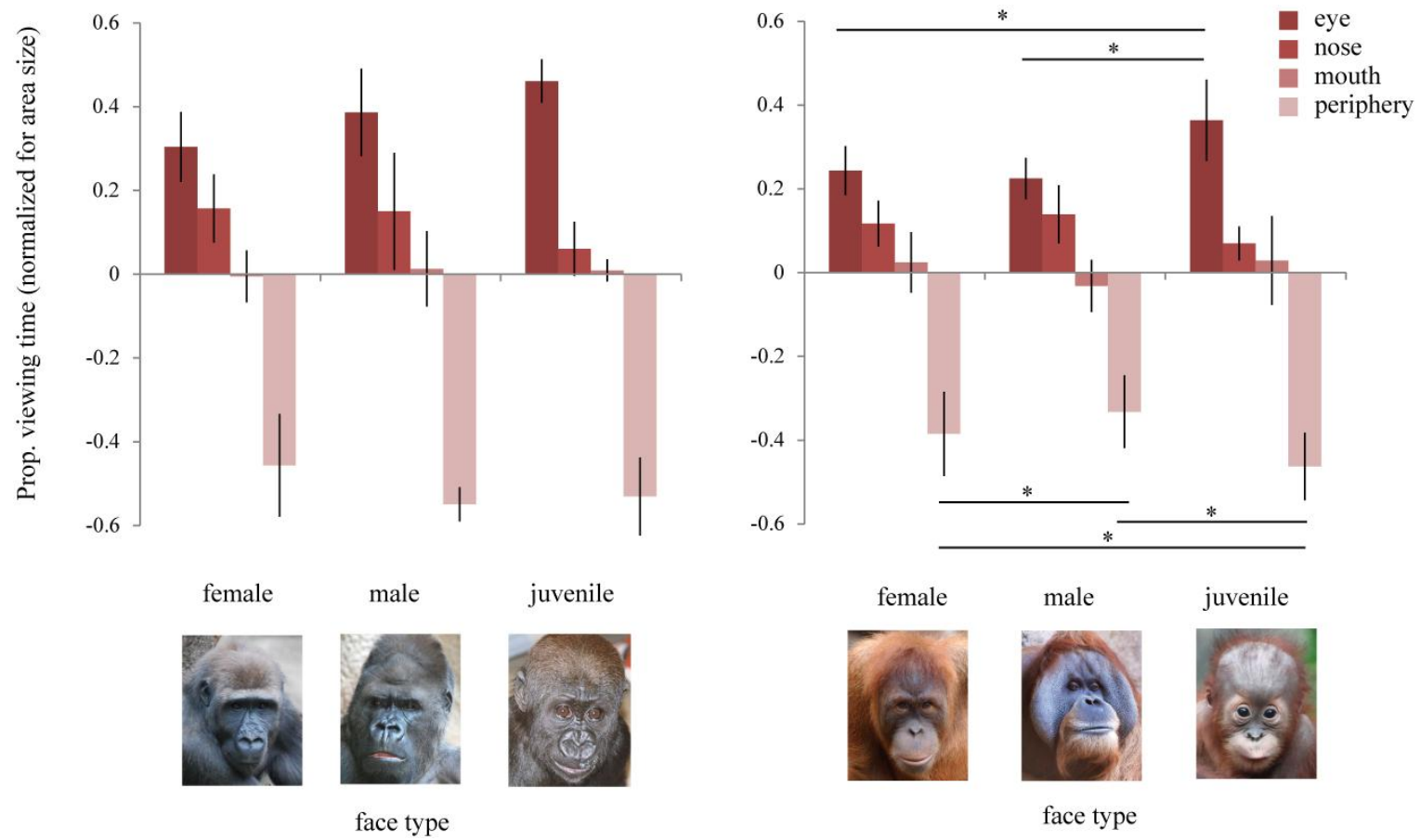


Figure 6

604



605

606 Figure 7

607

608 Supporting material

609

Table S1. Proportion of viewing time for each part of a scene (not normalized for size) in each species across studies.

Participant	Conspecific scene			Allospecific scene		
	Face	Body	Background	Face	Body	Background
Human (1)	0.7	0.21	0.08	0.66	0.28	0.05
Gorilla (1)	0.42	0.38	0.25	0.33	0.34	0.32
Orangutan (1)	0.45	0.29	0.19	0.34	0.43	0.22
Human (2)	0.57	0.21	0.21	0.72	0.18	0.09
Chimpanzee (2)	0.37	0.43	0.18	0.31	0.34	0.34

(1) This study
(2) Kano and Tomonaga (2009)

610

Table S2. Proportion of viewing time for each part of a face (not normalized for size) in each species across studies.

Participant	Conspecific face				Allospecific face			
	Eye (Eyeball)	Nose	Mouth	Periphery	Eye (Eyeball)	Nose	Mouth	Periphery
Human (1)	0.64 (0.23)	0.23	0.04	0.07	0.68 (0.26)	0.15	0.08	0.07
Gorilla (1)	0.48 (0.18)	0.2	0.13	0.17	0.41 (0.13)	0.07	0.18	0.31
Orangutan (1)	0.36 (0.15)	0.13	0.19	0.29	0.38 (0.12)	0.12	0.2	0.29
Human (2)	0.44 (0.27)	0.17	0.19	0.18	0.51 (0.28)	0.18	0.13	0.16
Chimpanzee (2)	0.37 (0.11)	0.16	0.25	0.2	0.28 (0.07)	0.24	0.15	0.31

(1) This study
(2) Kano and Tomonaga (2010)

Table S3. Average fixation duration (ms) for each AOI of whole-body/facial pictures.

	Human (1)	Gorilla (1)	Orangutan (1)	Human (2)	Chimpanzee (2)
Whole Body					
Face	821	375	356	1002	354
Body	299	212	286	313	207
Background	237	217	280	289	224
Whole Scene	607	255	318	739	259
Face					
Eye	410	293	357	484	257
Nose	303	250	337	415	275
Mouth	348	285	366	305	225
Periphery	287	232	318	272	221
Whole Face	374	282	366	464	255
(1) This study					
(2) Kano and Tomonaga (2009/2010)					

612

613